Three-Dimensional Molecular Modeling of Bovine Caseins: κ-Casein

T. F. KUMOSINSKI, E. M. BROWN, and H. M. FARRELL, JR.
US Department of Agriculture
Eastern Regional Research Center
Agricultural Research Service
600 East Mermaid Lane
Philadelphia, PA

ABSTRACT

Three-dimensional structures derived from X-ray crystallography are extremely important in elucidating relationships between structure and function for many proteins. However, not all proteins can be crystallized. The caseins of bovine milk are one class of noncrystallizable proteins. The complete primary and partial secondary structures of these proteins are known, but homologous proteins with known crystallographic structure are not available. In this report, sequence-based predictions of secondary structure were made and adjusted to conform with global secondary structures derived from Fourier transform infrared spectroscopy. With this information, a three-dimensional structure for κ-casein was constructed using molecular modeling computer programs. The constructed model contains two unstranded β-sheets; both are predominately hydrophobic and capable of forming quaternary structural interaction sites with α_{s1}-casein. This unrefined structure is in good agreement with much of the biochemical information available for κ-casein.

(Key words: casein structure, protein functionality, milk proteins)

Abbreviation key: FTIR = Fourier transform infrared.

INTRODUCTION

At the heart of the skim milk system is a unique biocolloid, the casein micelle. This colloidal complex is in dynamic equilibrium with its environment. Changes in the state of the casein micelle system occur during milk secretion and processing (13, 24). Because of its innate importance to the milk system, the casein micelle and its four major protein components have been studied extensively.

κ-Casein differs from other caseins in that it is soluble over a broad range of calcium ion concentrations (27). It was this calcium solubility that led Waugh and von Hippel (27), upon discovering the κ-fraction, to assign to it the role of casein micelle stabilization. It is also the κ-casein fraction that is most readily cleaved by chymosin (rennin) (16, 17); the resulting products are termed para-κ-casein and the macropeptide. It would appear that κ-casein is the key to micelle structure in that it stabilizes the calcium-insoluble α_{s1}- and β-caseins. Because reproducible crystals have not been obtained for κ-casein, crystallographic structures will probably not be realized. We have used molecular modeling computer programs to develop a proposed three-dimensional structure for κ-casein. This model is not an end in itself and should serve as a new starting point in the examination and exploration of casein structure and function.

MATERIALS AND METHODS

Predictions of Secondary Structures

Selection of appropriate conformational states for the individual amino acid residues was accomplished by comparing the results of sequence-based predictive techniques, primarily those of Chou and Fasman (6), Garnier et al. (14), and Cohen et al. (7, 8). Assignments of secondary structure (α-helix, β-sheet, or β-turn) for the amino acid sequences were made when either predicted by more than one method or strongly predicted by one and not predicted against by the others. In addition, because of the large number of proline residues in the caseins, proline-based turn predictions were made using the data of Benedetti et al. (4) and Ananthanarayanan et al. (2).

Received December 3, 1990.
Accepted February 25, 1991.

Molecular Modeling

The three-dimensional structure of κ-casein was approximated using molecular modeling methods with an Evans and Sutherland PS390[1] interactive computer graphics display driven by Sybyl (Tripos) (St. Louis, MO) software. The construction of this model was accomplished by drawing upon the predicted behavior of the polypeptide chain from its amino acid sequence and reconciling these predictions with spectroscopic data. In building the macromolecule, we used a library or dictionary of geometric parameters, i.e., bond lengths between specified atoms, bond angles, and van der Waals radii that are compatible with those found by X-ray crystallographic analysis. The molecular modeling software package contains this library or dictionary with average geometric parameters for each amino acid. Segments of the protein ranging in length from 60 to 80 amino acid residues were selected based upon secondary structural prediction. This represents the optimal length for such procedures, and the length was varied so as not to disrupt predicted runs of secondary structure elements. Each segment was then built amino acid by amino acid and assigned φ and ψ angles characteristic of the respective predicted structures for each residue. All ω angles were assigned the conventional *trans* configuration. In addition, aperiodic structures are in the extended rather than totally random configuration. The Sybyl subroutine "SCAN" was used, on the side chains only, to adjust torsional angles and relieve bad van der Waals contacts. The individual pieces were then joined together to produce the total polypeptide model and readjusted.

RESULTS AND DISCUSSION

Rationale for Generation of Secondary Structure Models and Three-Dimensional Models for Sequences with Multiple Proline Residues

Various methodologies for sequence-based secondary structural predictions are currently available (6, 7, 8, 14). These methods have been applied to the caseins (9, 18, 21). In our initial approaches, a sequence-based prediction was generated from each of three basis sets for κ-casein. The results were ambiguous; for some portions of the sequences, structures (α-helix, etc.) were consistently predicted, but in others they were not. The relatively high proline content of the caseins (Figure 1A) poses a problem because this residue, although occasionally found in both β-sheet and α-helical structures, is generally not favorable to either. In attempting to generate a consensus sequence-based prediction, we first focused upon solving the proline problem in a way commensurate with known behavior of this residue in proteins and model peptides (2, 4, 23).

The proline residues of the caseins are somewhat evenly dispersed throughout the sequences, ruling out both Type I and Type II polyproline structures (13, 24). It has been documented, however, in model peptides that proline frequently occupies the second position of either a four-residue β-turn or a three-residue γ-turn (2, 4, 23). Moreover, recent Fourier transform infrared (FTIR) spectroscopic analyses of the caseins have shown that up to 35% of the residues in caseins appear to be in β- or γ-turns [(5) and unpublished observations]. To assess the probability that proline residues in the caseins might be located in reverse turns, we examined the tetrapeptides containing proline in the second position for similarity with peptides known to form β- and γ-turns (2, 4). Most of the proline residues were predicted to be in a turn of some type. When the models were built, normal β-turns containing proline residues generally resulted in unfavorable van der Waals contacts with surrounding residues, leading us to assign most proline residues initially to the γ-turn conformation.

Two special cases are the sequences Pro-X-Pro (found 57-58-59 and 99-100-101) and Pro-Pro (residues 109-110 and 156-157) in κ-casein. Model peptides His-Pro-Pro-His and His-Pro-His-Pro-His were built to test the best φ and ψ angles to be inserted in these sequences. Because of its rigid ring structure, the φ and ψ angles for proline are almost always –70 and 70, respectively. Two proline residues side by side (X-Pro-Pro-X) result in a "kink" in the extended structure of κ-casein (Figure 2A). In the case of Pro-X-Pro, bad contacts and an unfavorable structure were observed

[1]Mention of brand or firm name does not constitute endorsement by USDA over products of a similar nature not mentioned.

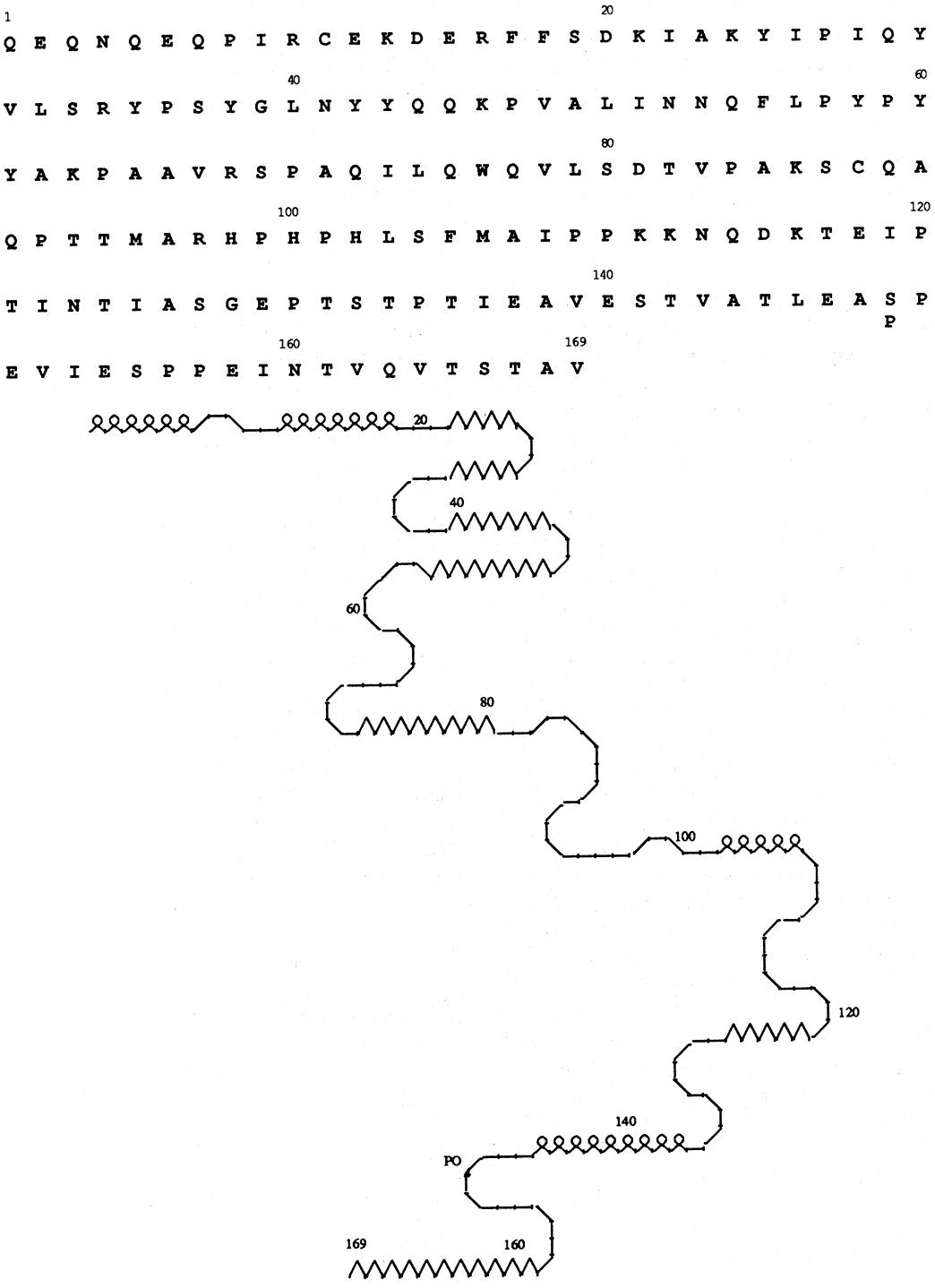

Figure 1. A) Sequence of κ-casein B. B) Summary of initial secondary structural assignments made for κ-casein.

TABLE 1. Comparison of adjusted sequence-based predictions with Fourier transform infrared (FTIR) spectroscopic data.

Sample	Helix	β-Structure	Turns	Unspecified
	(%)			
κ-Casein				
FTIR[1]	8	29	33	30
Model	16	27	37	20

[1]Preliminary data with 10% error [(5) and unpublished observations].

when angles greater than 70° were assigned (Figure 2B). When the center residue was assigned angles similar to proline, a more reasonable representation for the sequence (Pro-X-Pro), within an extended structure such as κ-casein, is shown in Figure 2C. This structure does not unduly constrain the polypeptide chain.

β-Turns, other than those based on proline, were predicted by the Chou-Fasman (6) and Cohen et al. (8) methods. The total number of amino acid residues predicted to be in turns was only slightly in excess of the total predicted by spectroscopic methods (Table 1).

In a similar fashion, "consensus" scores for α-helix and β-sheet were arrived at by choosing from those regions having the highest predicted probability of a given structure to yield values in accord with FTIR data [(5) and unpublished observations]. In this case, all residues previously assigned to proline-based turns were eliminated first from consideration in α-helical or extended β-structures. The net results of these calculations for κ-casein are compared with spectroscopic data in Table 1. Finally, all residues not included in these periodic structures were then considered to be in an extended aperiodic conformation. The initial conformational assignments for κ-casein are shown with its sequence in Figure 1. The κ-casein model resembles that of Loucheux-Lefebvre et al. (18), with the exception that the β- and γ-turns are included.

Rationale for Generation of Three-Dimensional Models

The secondary structural assignments that had been reconciled with FTIR spectroscopic data were used as a point of departure for the generation of three-dimensional models. Idealized ϕ and ψ angles assigned initially for each structural element are given in Table 2. Based upon naturally occurring breaks in predicted secondary structure (Figure 1B), protein segments consisting of 60 to 80 residues were constructed starting with the N-terminal end and proceeding toward the C-terminal end. Each segment was examined visually for "bad contacts" and adjusted for unfavorable van der Waals interactions of the side chains. The individual pieces were then joined to produce the total polypeptide, and torsional angles were again adjusted to remove unfavorable van der Waals contacts among the side chains. The initial backbone conformation was maintained throughout this procedure. This procedure is analogous to the mechanism whereby the protein is synthesized in vivo from N → C terminal and presumably folds after insertion into the lumen of the endoplasmic reticulum (13).

TABLE 2. Dihedral (ϕψ) angles assigned to specific conformational states in the initial structures for κ-casein.

	Angle assignments	
Structures	ϕ	ψ
α-Helix	−58	−47
β-Sheet	−139	135
β-Turn		
N	−139	135
N + 1	−60	−30
N + 2	−90	0
N + 3	−139	135
γ-Turn		
N	135	−69
N + 1	−75	59
N + 2	81	−126

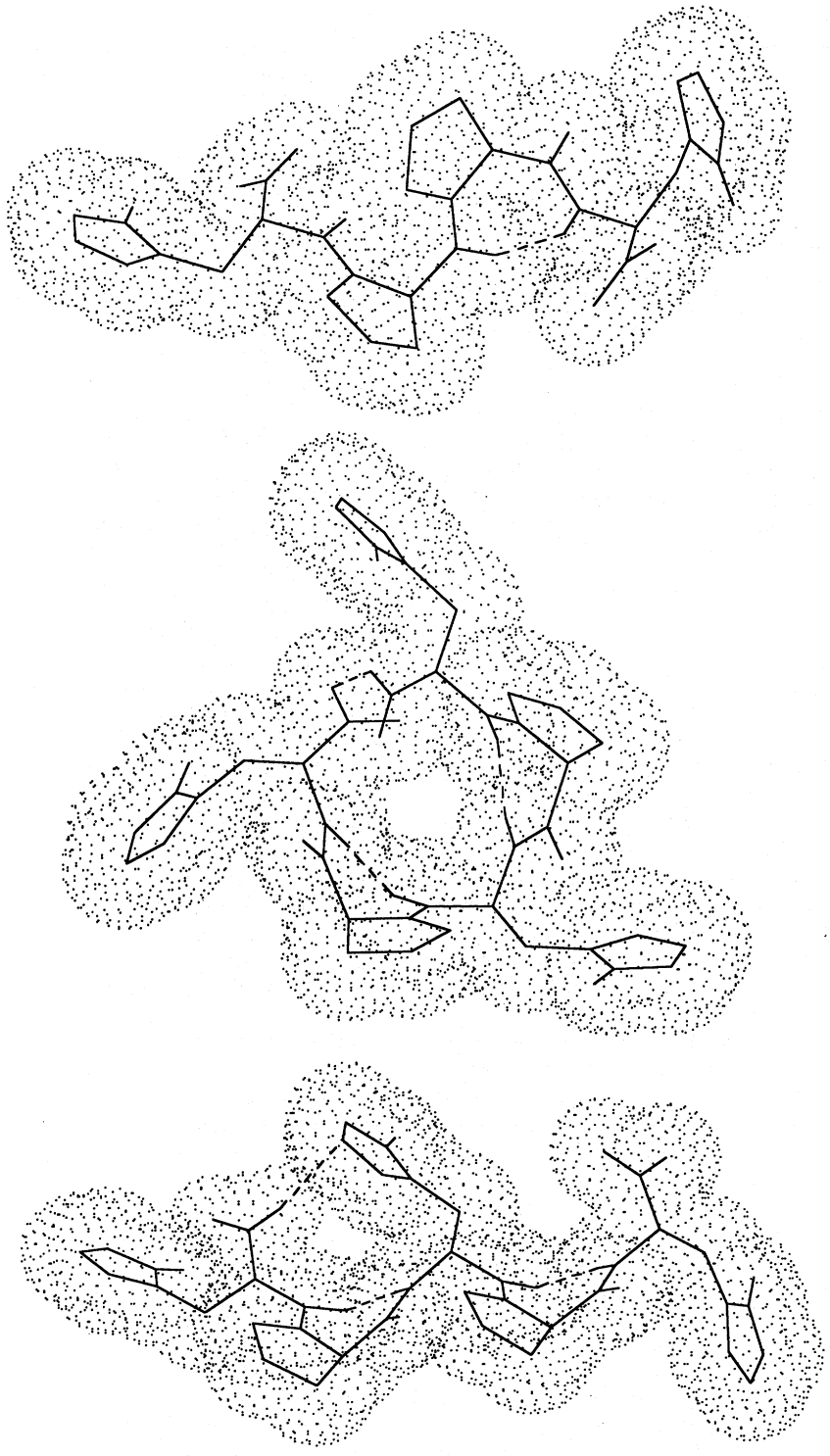

Figure 2. A) Model peptide for X-Pro-Pro-X sequence demonstrating an extended structure (X = His). Model peptides for the structure X-Pro-X-Pro-X: B) shows a highly constrained configuration; the second C) is the more likely structure used here (X = His).

Three-Dimensional Molecular Model of κ-Casein

The computer model generated as described for κ-casein is shown in Figure 3. In descriptive terms the protein can be thought of as being represented by a "horse and rider". The amino-terminal 110 to 120 residues represent the "horse", and the carboxy-terminal portion represents the "rider." Two distinct legs are seen in the "horse" portion of the model. These legs are generated by β-sheet regions comprising residues 20 to 25; 29 to 34; 39 to 45; and 49 to 55, which are connected by γ- or β-turns. The overall dimensions of the κ-casein monomer predicted by this model are 8.0 × 6.2 × 5.8 nm. In the following section, we will attempt to reconcile some known features of the chemistry of κ-casein with this molecular model. To aid in this discussion, the backbone-only model with prolines indicated is given in Figure 4A, along with a stereo view (Figure 4B).

Chemistry of κ-Casein and the Three-Dimensional Model

Chymosin (Rennin) Hydrolysis. The action of chymosin on the casein micelle is primarily the hydrolysis of the highly sensitive Phe-Met peptide bond (residues 105 to 106) of κ-casein. Sequence data (19) show this bond to be in a proline-rich region of the molecule between Pro-His-Pro (residues 99-101) and Pro-Pro (residues 109-100), which perhaps accounts for the high susceptibility of this specific bond to chymosin. From studies of model peptides, it has been suggested that the residues lying between Pro-101 and Pro-109 occur in a β-sheet structure (19). Using other predictive methods, there is an equal chance that these residues could be in an α-helical conformation. In either case, the Pro-X-Pro and Pro-Pro residues cause the formation of a kink, which neatly presents the otherwise hydrophobic Phe-Met on the surface of the molecule. In our model, we show the α-helix that represents the "horse's bit" in the descriptive "horse and rider" model. From model building considerations, this sequence represents the minimum number of residues for a stable α-helix. Cleavage of the Phe-Met (105-106) bond would render the helical conformation untenable, the helix would unwind, and a considerable amount of configurational entropy would be added to the hydrolysis reaction. The change from α-helix to extended conformation, although significant, would be lesser in its energy contribution. In either case, helix or sheet, the Pro-Pro and Pro-X-Pro turns present this segment on the surface and make it readily accessible to chymosin. Any model for casein micelle structure must in some way account for this feature of κ-casein.

Hydrophobic Interactions. Earlier speculation (15) that κ-casein might be a linear amphiphile seems to be only partially true. The amino-terminal fifth of the molecule has a relatively high charge frequency (frequency = .34); however, the net charge is zero, and this part of the protein, although relatively hydrophobic, is somewhat exposed (Figure 4). Residues 20 to 68 represent an exceptionally hydrophobic area with almost no charge. It is precisely within this region that the majority of the residues found in the "legged" structures of the κ-casein molecule occur. These nonstranded, highly hydrophobic β-sheets make ideal sites for sheet by sheet interactions with other κ-casein molecules or with hydrophobic domains of α_s- and β-caseins. Indeed, the concentration-dependent reaction profile of the reduced form of purified κ-casein can be fitted with a model for polymerization at a critical micelle concentration of .05% (26). Several investigators (1, 10, 11, 22) have noted that β- and κ-caseins diffuse out of the casein micelle at low temperatures. As the temperature decreases, hydrophobic stabilization energy decreases, and κ-casein is able to dissociate from the micelle. Finally, the importance of the hydrophobic region in casein-casein interactions can be supported by the research of Woychik (28). Of nine tyrosine residues in κ-casein, seven are located between residues 35 and 68; nitration of seven tyrosines in κ-casein severely inhibited its ability to stabilize α_{s1}-casein (28). As can be seen in Figure 3, the "legged" structures are constituted from this region and contain seven tyrosine residues.

Sulfhydryl and Disulfide Interactions. κ-Casein contains two cysteine residues. Whether these can form intra- or intermolecular disulfide bonds and the effects of such bonding on micelle stabilization have not been clearly established. The occurrence of free sulfhydryl groups in the milk protein complex has been reported by Beeby (3), but not by others (16). Swaisgood et al. (25) reported significant S-S crosslinking in purified κ-casein; however, Woychik et al. (29) demon-

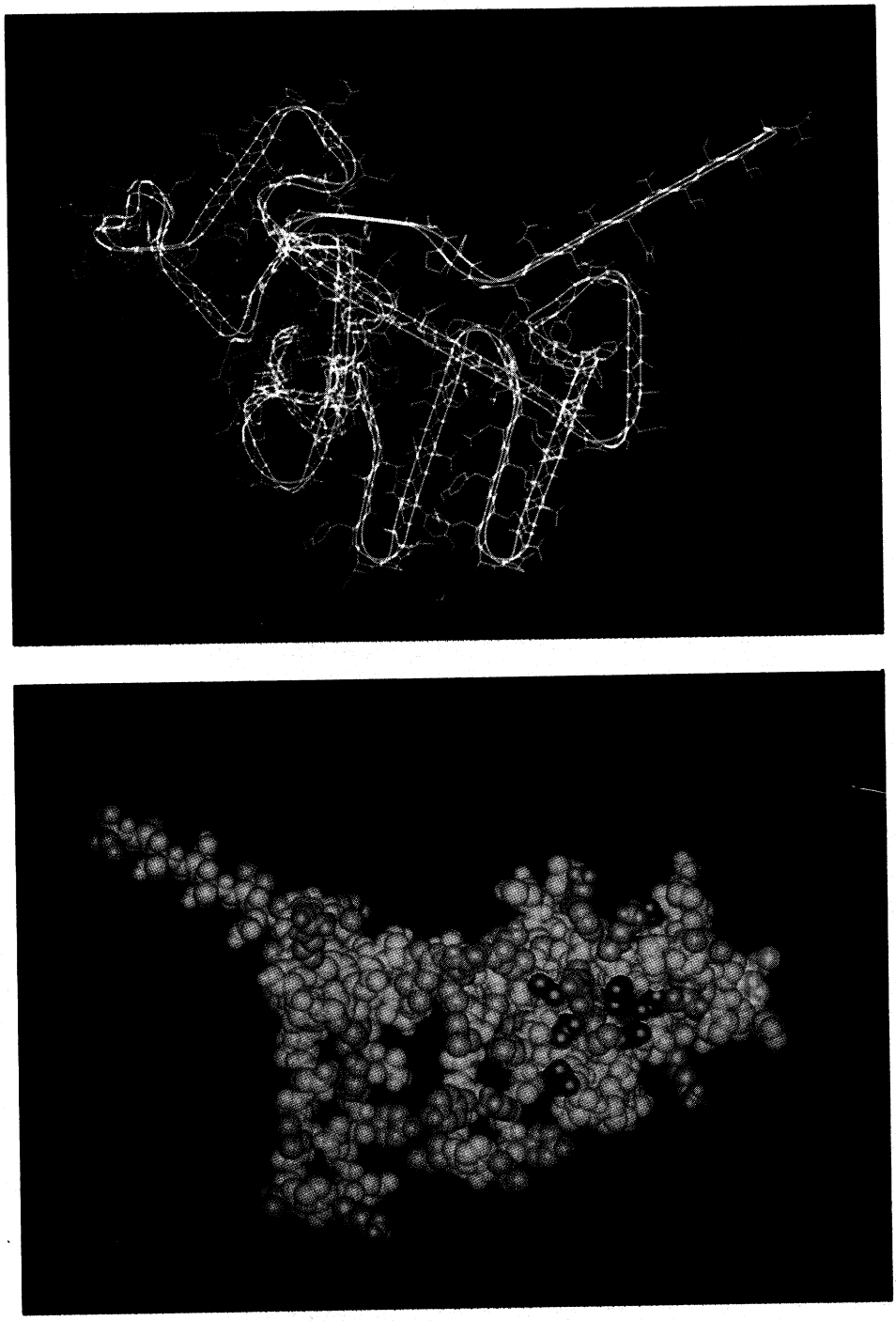

Figure 3. A) Three dimensional molecular model of κ-casein. The peptide backbone is colored cyan, hydrophobic side chains green, acidic side chains red, and basic side chains purple. B) Space filling model; Phe-Met (105-106) in orange.

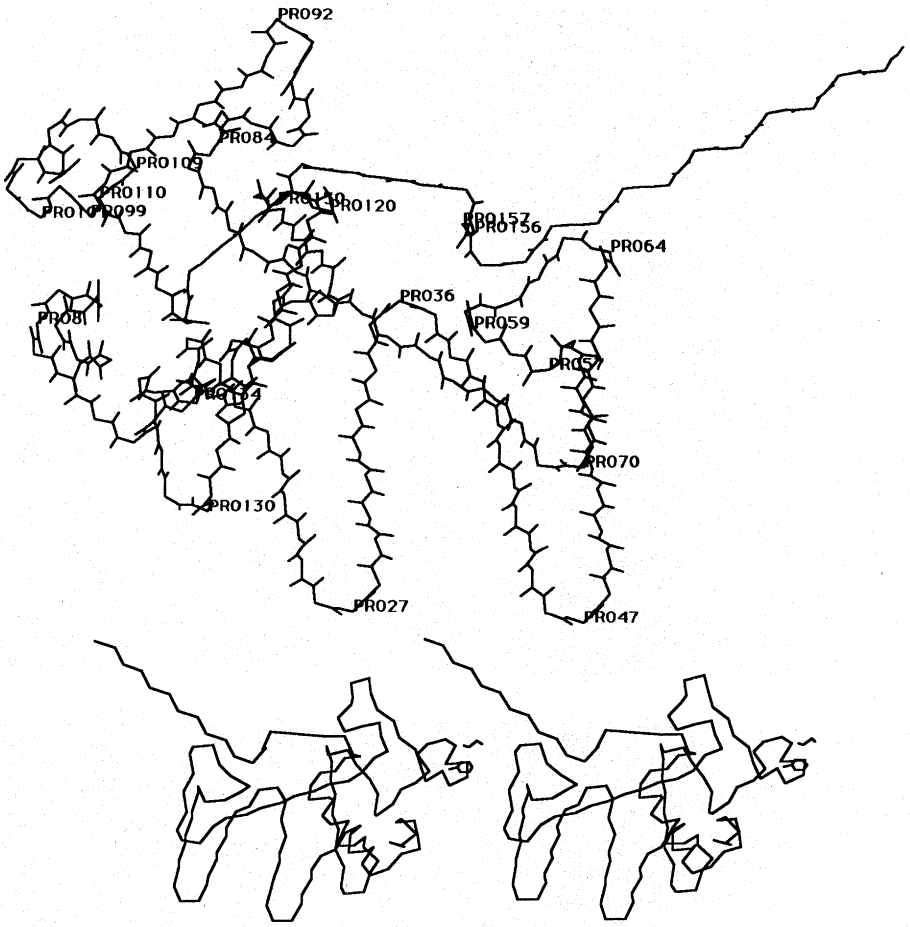

Figure 4. A) Chain trace of κ-casein; prolines indicated. B) Stereo view of the three-dimensional molecular model of κ-casein, showing the backbone without side chains, except Phe-105 and Met-106.

strated that reduced and alkylated κ-casein stabilized α_{s1}-casein against calcium precipitation as well as native κ-casein. Pepper and Farrell (20) found that in soluble whole casein, in the absence of Ca^{2+}, κ-casein occurred as a high molecular weight polydisperse complex. At low protein concentrations, this complex could be separated from the other caseins by size exclusion gel chromatography in the absence of urea. The addition of reducing agents converted the κ-casein to the sulfhydryl-form, which exhibited concentration-dependent associations, both with itself and with other caseins. Thus, although κ-casein represents only 13% of the casein, many κ-casein molecules must be somewhat contiguous in the micelle in order to form these disulfide-linked aggregates. It is interesting to note that Cys-11 is located between two segments predicted to be in α-helical conformations and is found on the "rider's" left, and Cys-88 is located in a predicted β-turn on the opposite side. In our model, both of these residues are located near the surface of the molecule and are directed away from each other at a distance of over 33 Å. This could account for the ability of the κ-casein molecule to form the interchain disulfide bonded polymers as discussed.

Sites for Glycosylation and Phosphorylation of κ-Casein. Of the major components of the casein complex, only κ-casein can be glycosylated. Nearly all of the carbohydrate, as well as the phosphate associated with κ-casein, is bound to the macropeptide (12), which is the highly soluble portion liberated by chymosin

hydrolysis. The major site for glycosylation, Thr-133, is found in our model on the back of the "horse" and is on a β-turn. The sites of phosphorylation, Ser-149 and Thr-145, are on the back portion and are found in β-turns as well.

Refinement of Casein Structures

It must be stressed that the structures presented in this work represent preliminary models. They have been partially refined and poor van der Waals side chain contacts removed, but they are not energy minimized structures. Thus, electrostatic interactions, hydrogen bond formation, and backbone interactions are not taken into account. The refinement of the entire structure through the use of tools such as the Kollman force field will be the thrust of future work. However, it is imperative to note that even this unrefined molecular modeling technique (when combined with predicted secondary structures and spectroscopic results) can reveal important relationships between structure and function. This information can in turn be used for designing new functional proteins using site-directed mutagenesis.

REFERENCES

1 Ali, A. E., A. T. Andrews, and G. C. Cheeseman. 1980. Influence of storage of milk on casein distribution between micellar and soluble phases and its relationship to cheese making properties. J. Dairy Res. 47: 371.

2 Ananthanarayanan, V. S., S. K. Brahmachari, and N. Pattabiraman. 1984. Proline-containing β-turns in peptides and proteins: analysis of structural data on globular proteins. Arch. Biochem. Biophys. 232:482.

3 Beeby, R. 1964. The presence of sulfhydryl groups in κ-casein. Biochim. Biophys. Acta 82:418.

4 Benedetti, E., A. Bavoso, B. D. Blasio, V. Pavone, C. Pedone, C. Toniolo, and G. M. Bonora. 1983. Solid-state geometry and conformation of linear, diastereoisometric oligoprolines. Biopolymers 22:305.

5 Byler, D. M., and H. M. Farrell, Jr. 1989. Infrared spectroscopic evidence for calcium ion interaction with carboxylate groups of casein. J. Dairy Sci. 72: 1719.

6 Chou, P. Y., and G. D. Fasman. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47:45.

7 Cohen, F. E., R. M. Abarbanel, I. D. Kuntz, and R. J. Fletterick. 1983. Secondary structure assignment for α/β proteins by combinatorial approach. Biochemistry 22:4894.

8 Cohen, F. E., R. M. Abarbanel, I. D. Kuntz, and R. J. Fletterick. 1986. Turn prediction in proteins using a pattern-matching approach. Biochemistry 25:266.

9 Creamer, L. K., T. Richardson, and D.A.D. Parry. 1981. Secondary structure of bovine α_{s1}- and β-casein in solution. Arch. Biochem. Biophys. 211:689.

10 Davies, D. T., and A.J.R. Law. 1983. Variation in the protein composition of bovine casein micelles and serum caseins in relation to micellar size and milk temperature. J. Dairy Res. 50:67.

11 Downey, W. K., and R. F. Murphy. 1970. The temperature dependent dissociation of β-casein from bovine casein micelles and complexes. J. Dairy Res. 37:361.

12 Eigel, W. N., J. E. Butler, C. A. Ernstrom, H. M. Farrell, Jr., V. R. Harwalker, R. Jenness, and R. McL. Whitney. 1984. Nomenclature of the proteins of cow's milk: fifth revision. J. Dairy Sci. 67:1599.

13 Farrell, H. M., Jr. 1988. Physical equilibria: proteins. Page 461 *in* Fundamentals of dairy chemistry. 3rd ed. N. Wong, ed. Van Nostrand Reinhold, New York, NY.

14 Garnier, J., D. J. Osguthorpe, and B. Robson. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. 120:97.

15 Hill, R. J., and R. G. Wake. 1969. Amphiphile nature of κ-casein as the basis for its micelle stabilizing property. Nature (Lond.) 221:635.

16 Jolles, P., C. Alais, and J. Jolles. 1962. Amino acid composition of κ-casein and terminal amino acids of κ- and para κ-casein. Arch. Biochem. Biophys. 98:56.

17 Kalan, E. B., and J. H. Woychik. 1965. Action of rennin on κ-casein, the amino acid composition of para κ-casein. J. Dairy Sci. 48:1423.

18 Loucheux-Lefebvre, M.-H., J.-P. Aubert, and P. Jolles. 1978. Prediction of the conformation of cow and sheep κ-casein. Biophys. J. 23:323.

19 Mercier, J. C., G. Brignon, and B. Ribadeau-Dumas. 1973. Primary structure of κ-casein B bovine. Complete sequence. Eur. J. Biochem. 35:222.

20 Pepper, L., and H. M. Farrell, Jr. 1982. Interactions leading to formation of casein submicelles. J. Dairy Sci. 65:2259.

21 Raap, J., K.E.T. Kerling, H. J. Vreeman, and S. Visser. 1983. Peptide substrates for chymosin (rennin): conformation studies of κ-casein and related oligopeptides by circular dichroism and structural predictions. Arch. Biochem. Biophys. 221:117.

22 Rose, D. 1968. Relation between micellar and serum casein in bovine milk. J. Dairy Sci. 51:1897.

23 Rose, G. D., L. M. Gierasch, and J. A. Smith. 1985. Turns in peptides and proteins. Adv. Prot. Chem. 37:1.

24 Schmidt, D. G. 1982. Associations of caseins and casein micelle structure. Page 61 *in* Developments in dairy chemistry–1. P. F. Fox, ed. Appl. Sci. Publ. Ltd., Essex, Engl.

25 Swaisgood, H. E., J. R. Brunner, and H. A. Lillevik. 1964. Physical parameters of κ-casein from cow's milk. Biochemistry 3:1616.

26 Vreeman, H. J., J. A. Brinkhuis, and C. A. Van Der Spek. 1981. Some association properties of bovine SH-κ-casein. Biophys. Chem. 19:185.

27 Waugh, D. F., and P. H. Von Hippel. 1956. κ-Casein and its stabilization of casein micelles. J. Am. Chem. Soc. 78:4576.

28 Woychik, J. H., and M. V. Wondolowski. 1969. Nitration of tyrosyl residues in κ- and α_{s1}-caseins. J. Dairy Sci. 52:1669.

29 Woychik, J. H., E. B. Kalan, and M. E. Noelken. 1966. Chromatographic isolation and partial characterization of reduced κ-casein components. Biochemistry 5:2276.